An Ontology-Based Data Model to Create Virtual Training Environments for Construction Safety Using BIM and Digital Twins

Kilian Speiser, Jochen Teizer
Technical University of Denmark (DTU), Denmark
kilsp@dtu.dk

Abstract. Accident rates in the construction industry remain high. Recent research introduced serious games based on extended reality and proved increased safety awareness among the trainees. However, creating such training scenarios requires resources across many domains. This work aims to validate that a Digital Twin for Construction (DTC) provides the necessary data to generate a virtual training environment automatically. We propose a novel ontology-based data model integrating a serious game into a DTC. The data model includes the project intent information, safety regulations, and status knowledge to update the learning environment dynamically. A prototype is developed in Unity based on two test projects (building and infrastructure) that contain several typical construction hazards. The case study indicates the data model's viability for generating a virtual environment with little means based on the DTC. This work further suggests value once the underlying data sources for the model are connected and of high quality.

1. Introduction

The construction industry is one of the most hazardous industries worldwide. In the United States alone, an average of 980 workers in the construction industry die every year due to occupational accidents on construction sites (CPWR, 2022). Compared to other working environments, construction and extraction workers have one of the deadliest professions (BLS, 2022). To address this problem, many organisations invest heavily in safety training programs to reduce the number of accidents and fatalities by improving safety knowledge and awareness. Traditional training methods, such as classroom instruction and on-the-job training, have limitations that can reduce their effectiveness. Particularly for safety training, the method must not expose the trainees to hazards. Recent research has introduced extended reality (XR) games as a potential solution for safety training in construction. Immersive technologies in XR enhance virtual training environments. Studies have shown that these technologies can improve the effectiveness of construction safety training by providing trainees with realistic simulations of construction tasks and hazards (Sacks et al., 2013; Hilfert et al., 2016; Nykänen et al., 2020). Virtual Training Environments (VTEs) allow for providing personalised feedback based on collected run-time data (Golovina et al. 2019a; Golovina and Teizer, 2022). Personalised feedback is crucial for achieving optimal training outcomes (Pianta et al., 2012). Nevertheless, creating these games is time-consuming and lacks realism (Bükrü et al., 2020; Jacobsen et al., 2022; Wolf et al., 2022; Wolf and Teizer, 2022). Another issue is that trainees can suffer virtual reality sickness during or after immersive experiences. Even with the evolvement of better hardware, one out of three users experience discomfort, and at least 5% encounter severe symptoms (Stanney et al., 2020).

Digital Twins (DT) digitally represent physical assets, systems, and processes (physical twin) that can monitor, analyse, and optimise their performance in real-time. The digital twin for construction (DTC) remains somewhat undefined. According to Sacks et al., 2020, federated models that depict the as-planned and as-designed states of a constructed asset do not qualify as a DT. Instead, a DTC requires the frequently updated as-built and as-performed states to generate insights into construction processes, expenses, and materials. This data is vital as a

DTC compared to Building Information Modelling (BIM) does not only define the flow of information but also specifies a comprehensive approach to construction planning and management. In the context of this work, we define the DTC as a comprehensive digital replica of a building project, including information about the building's geometry, structure, systems, and processes. By integrating data from different sources such as BIM models, Internet of Things (IoT) sensors, and other relevant information, DTCs can provide a holistic view of the construction project and enable more accurate and efficient decision-making throughout its lifecycle. With the increasing adoption of DTCs in the industry, the interest in leveraging them for safety training grows. Teizer et al. (2022) propose a Digital Twin for Construction Safety (DTCS), and Harichandran et al. (2021) conceptually describe a dynamic, serious game for construction safety as part of a DTCS. Nevertheless, no previous study has verified this framework.

Multiple studies with trainees in a VTE performing construction work tasks proved effective ways of simulating the exposure of workers to hazards without exposing them to actual hazards. However, the creation of virtual training scenarios requires extensive resources. Thus, the main hypothesis of this work is that a DTC provides data sources that ease the creation of VTEs based on BIM. Hence, this study provides a data model that can generate and update VTEs for construction safety using the information from a DTC. After introducing the applied method to develop the data model, Section 3 describes the ontology-based model, which utilises existing schemas for the built environment, and extends these with crucial elements regarding construction safety and personalised feedback. To verify the effectiveness of the data model, Section 4 describes a prototype using the game engine Unity, which we tested using two real projects from the education and infrastructure construction sectors: a college building and a railway project.

2. Model Development

The research method for developing the data model includes four steps: specification, knowledge acquisition, implementation, and validation. Despite Zheng et al. (2021) introducing this approach to specify ontologies, Schlenger et al. (2022) proved its applicability for data models. First, the data model's high-level goals, objectives, and requirements were specified. There are existing models for civil engineering domains, such as the Building Topology Ontology (BOT) or IFC. Bernardos et al. (2022) present an ontology for a DTC, and Schlenger et al. (2022) propose a process-oriented data model for a DTC. Johansen et al. (2023) extend existing ontologies with a hazard ontology for construction safety. Still, no ontology or model considers the aspects required for generating VTEs for construction safety. Thus, the goal is to fill this gap with a simple data model that is extendable for domain-specific needs.

Similar to the framework from Ackoff (1989), the authors opted to structure the overall data into three layers. These layers - data, information, and knowledge - can be visualised as horizontal layers of a pyramid, where the value and conciseness of the data increase from bottom to top. In the context of virtual training for construction safety, the data layer represents raw data such as trajectories of players or locations of hazards and equipment or BIM models. An extensive data analysis derives the information layer. For instance, close calls can be detected and assessed based on the trajectories. Further evaluation leads to knowledge acquisition, such as worker safety awareness and safety design malfunctioning.

Simple data structures suffice for the raw data layer. Thus, existing ontologies require few adjustments. However, the knowledge layer requires reasoning on lower levels. Therefore, the information layer must have an adequate representation. The authors defined user stories for this work to specify the main stakeholders and identify crucial relationships. The main high-level requirements are:

- Automatically generate training scenes with a minimum of user interaction
- Include the most recent BIM model from DTC and add it to the training scene
- Retrieve the construction schedule from DTC, including tasks, resources, dates, and time
- Provide personalised feedback based on run-time data collection

After setting high-level objectives and requirements, a literature review identified existing ontologies and data models. While creating the schema, we carefully analysed overlaps and differences with existing models to ensure alignment with the current state-of-the-art.

3. Proposed Ontology-Based Data Model to Generate Training Environments

Figure 1 depicts the developed data model. The figure utilises the Unified Modeling Language (UML) to illustrate the design (Grady et al., 2005). The components in yellow refer to the industry standard IFC. While the green components are specific elements for the virtual training, the other components are general entities from the DTC. The five different types of relationships also refer to UML.

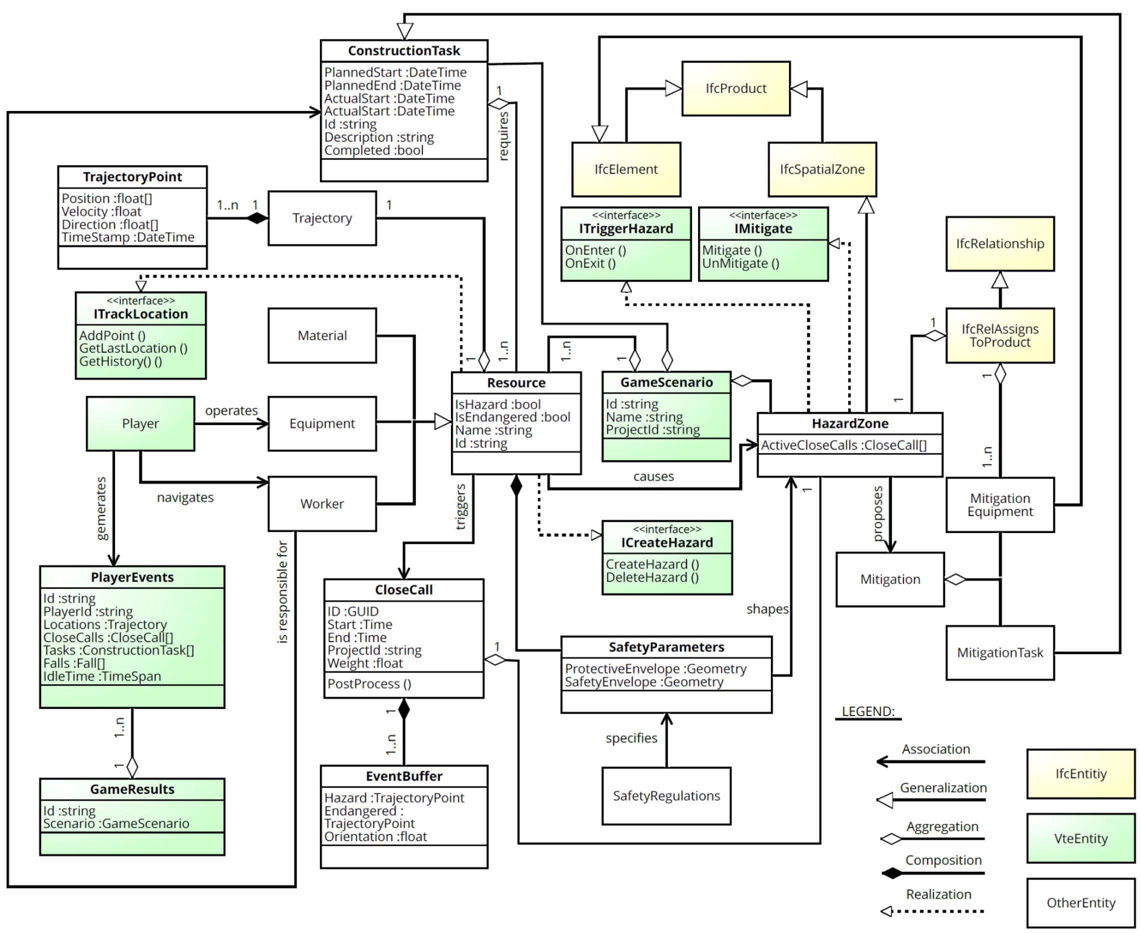

Figure 1: UML diagram of the proposed data model.

The *GameScenario* defines high-level information for the trainer and trainee. It extracts important data from construction tasks the trainer wants trainees to practise. A *GameScenario* may implement multiple construction tasks to allow for the simultaneous training of different actors. Each *ConstructionTask* has a responsible worker, a planned start and end date, an actual start and end date, and the required resources. The training environment is created based on as-planned 3D models and training tasks (*ConstructionTask*) from the construction schedule provided by the DTC. With the task as the initial input, the training scene requests the status of the as-planned construction model at the planned start time of the task from the DTC and adds all the active elements to the training scene. This includes, among others, physical objects, relationships between objects, properties, and geographic references. The construction schedule must connect the construction elements with construction tasks.

In this context, we specify requirements for the BIM model referring to the IFC schema. However, this work does not present a holistic model view definition. Li et al. (2022) propose a SafeConDM model describing a domain model for safety design, and Johansen et al. (2023) describe a hazard ontology. This work assumes that the DTC provides such a "safe" 4D BIM model, including the construction site layout and safety design. A successful implementation requires the provided BIM model to define hazard zones and mitigation equipment.

This research defines a *HazardZone* as space workers can only enter if they are specifically permitted to. Hazard zones must be part of the BIM model and are entities of the type IfcSpatialZone (Johansen et al., 2023). workers collide with the zone, they put themselves in danger and trigger a close call (class: *CloseCall*). Thus, the *HazardZone* realises the interface *ITriggerHazard* which defines methods for entering and leaving the zone. A vital component of this model is that *HazardZones* propose a *Mitigation* in the form of *MitigationEquipment*. For example, a safety guardrail (*MitigationEquipment*) mitigates an unprotected leading edge, or signs and barriers mitigate an electrocution hazard. The interface *IMitigate* provides methods for adding and removing this equipment in the training scene. The mitigation equipment must be removed automatically based on user input to expose the workers to hazards. For instance, the trainer wants to remove 5% of all guardrails for fall hazards. This information is part of the game scenario. To allow these settings, the data model utilises the *IfcRelAssignsToProduct* association within the IFC schema for relating the mitigation to a hazard.

Each hazard has a hazard type assigned. This model includes the four leading causes of fatal injuries in construction: falls to a lower level, electrocution, caught-in-between, and struck-by (Bureau of Labor Statistics, 2022). Typically, falls from heights and electrocution are static zones. At the same time, struck-by and caught-in-between hazards are caused by dynamically moving construction equipment such as crane loads or heavy machines.

Resource is an abstract class generalising *Worker*, *Material*, and *Equipment*. Resources have a trajectory providing the current and historical location and defining whether a resource is a hazard or endangered. For an endangered resource (human worker), the *SafetyParameters* geometrically describe a protective envelope and a safety envelope. However, the safety parameters shape the hazard zone if a resource is a hazard. As the resource may move, hazard zones can be dynamic. The interface *ITrackLocation* provides the required methods. Figure 2 describes this concept for a human worker and a machine. For the worker, the safety envelope is a bounding box around the body, and a cylinder represents the protective envelope. An oriented bounding box around a hazard with an offset of 1m on each side represents the safety envelope for the equipment. Referring to ISO 5006, a sphere with a radius of 12 meters specifies the protective envelope of earth-moving machinery (Golovina et al., 2019b; Golovina et al., 2021). Material can also depict a hazard zone, i.e., a loose cable on the ground may cause a trip.

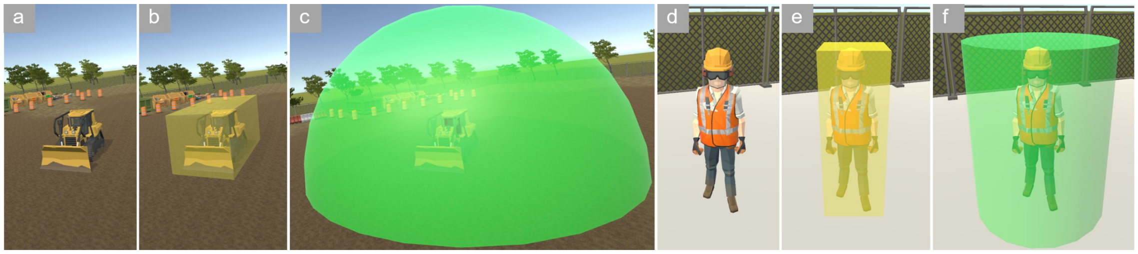

Figure 2: Safety parameters of a dozer and a pedestrian worker.

According to the Occupational Safety and Health Administration (2016), a *close call* is an incident where an accident almost occurs but is prevented at the last moment. A close call can involve the risk of death, injury, or illness, but chance or timely intervention avoids these outcomes. Close calls are significant because they indicate that workplace hazards are not adequately controlled and may require further attention to prevent future accidents. Within this study, a "close call" refers to a situation in which the protective envelope of an endangered resource collides with a hazardous zone (see Figure 3). A close call terminates once the endangered resource has left the hazard zone. Thus, every close call includes two actors: The endangered resource and the hazard. The buffer events provide a more comprehensive understanding of these situations. Each *EventBuffer* includes the timestamp, distance between the actors, the location velocity, and the direction of both actors. This information is crucial for accurately assessing the severity of the close call. Figure 3c illustrates the recorded data for a close call.

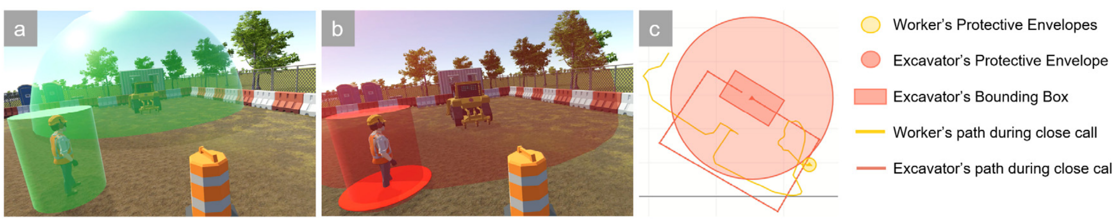

Figure 3: (a) Dozer representing hazard, (b) a collision of the safety envelopes, and (c) the recorded trajectory data related to the close call.

The player navigates an avatar or operates a machine in the virtual environment. During the training, the trajectories of resources are recorded. The trajectories are essential for evaluating the player's performance and retrieving valuable behavioural patterns. For that reason, each resource implements the *ITrackLocation* interface that continuously posts the location, timestamp, resource-id, and game-id to the *GameResults*. These results contain information about close calls and trajectories and function as a base for generating knowledge that can support decision-makers. For instance, algorithms can evaluate the player's performance based on KPIs. The proposed data model suggests a few indicators, such as the walked distance, the close call frequency, close call weights, or idle time.

4. Implementing a Prototype

To verify the proposed model, we developed a framework of a DTC integrating a VTE environment for construction safety. Figure 4 illustrates the architecture of the prototype. The cloud-based DTC provides the required information through an API that implements endpoints to request the BIM model, the construction schedule, and resources. The training environment itself was developed in Unity and hosted in a web application using WebGL. The VTE collects data at run-time and streams it to the database. The web application comprises two other components to provide personalised feedback to the trainees and summarised training results to the trainer. This prototype was tested for two projects: A college building and a railway project. The following sections present the results of the tests.

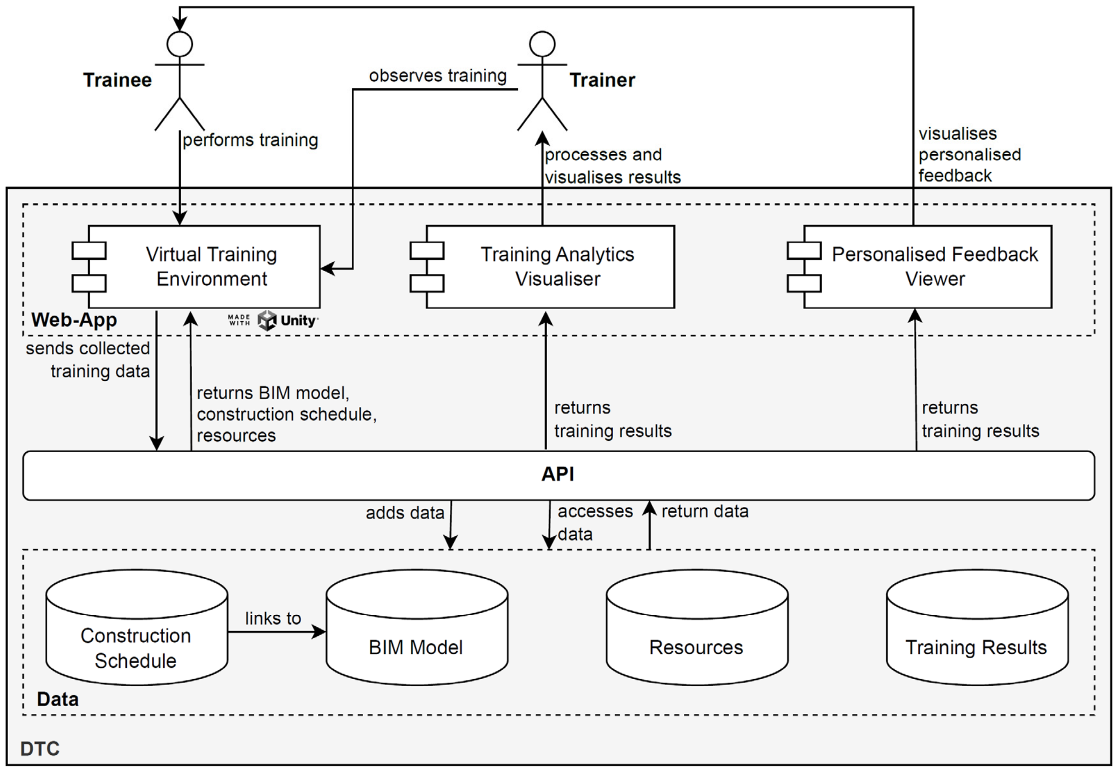

Figure 4: System architecture for the prototype.

4.1 Generating the Training Scene in Unity

The main objective of this work was to generate the training environments with little manual effort independently of the project or domain and collect data during run-time. The VTE requests an IFC file from the DTC. To allow importing the model into Unity at run-time, *IfcConvert* generates an OBJ, MTL, and XML file (IfcOpenShell, 2023). These files are loaded in Unity using a modified version of the *IfcImporter* asset (Arcventure, 2023). The import results in an object with the hierarchy of the original IFC file. In the next step, the VTE requests the construction schedule from the DTC and filters the active elements at the given date. The scene at this stage contains mitigated hazards such as safety guardrails. Thus, this equipment is removed based on the information from the training scenario.

Figure 5 indicates that the generation of the training scene was successful. The script added the BIM model at the given timestamp for both projects. Minor issues relate to the fact that the application IfcConvert is at the beta stage and not officially certified for IFC. In buildings, unprotected leading edges and crane loads are typical hazards. Thus, the training scenario includes three fall hazards and a crane load. For the infrastructure project, the training scenario implements one unprotected leading edge, an excavator, and a dozer.

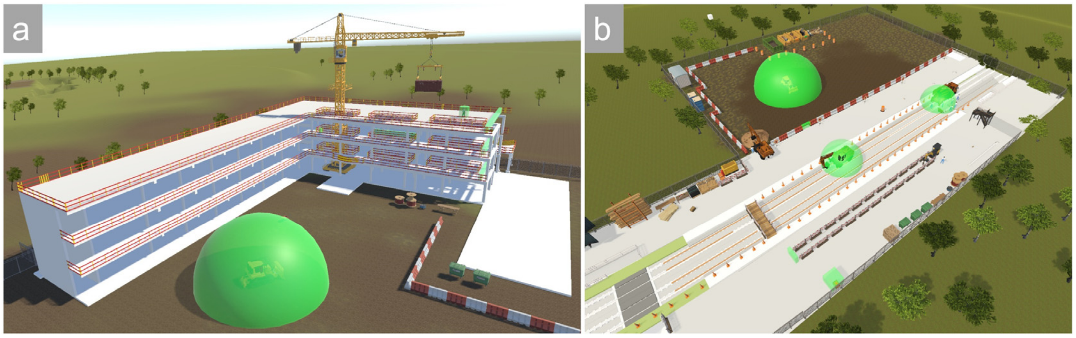

Figure 5: Virtual Training Environments (VTE) for the (a) building and (b) infrastructure project.

4.2 Validation Examples

The following examples validate the correct generation of resources. The player navigates too close to an unprotected leading edge and triggers a close call. The hazard zone implements the *ITriggerHazard* interface, which triggers upon a collision of the avatar's PE. Once the avatar leaves the zone, the close call is post-processed. Figure 6 visualises the results and indicates the model's viability for this application. The velocity weight is high as the player approaches the hazard with high speed. The orientation, duration, and deviation weights are low, implying that the player approached the hazard at an acute angle and left it immediately without detours.

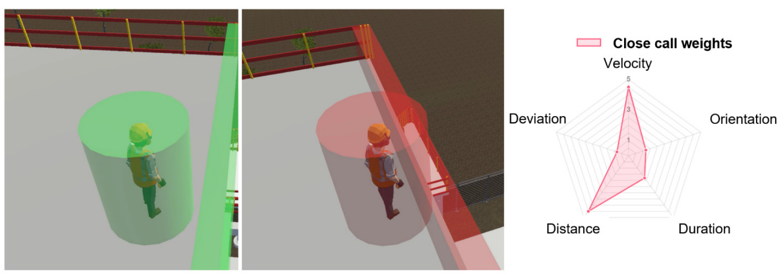

Figure 6: The worker triggers a close call colliding with the fall hazard zone.

The prototype for the infrastructure project contains two dynamic hazards: a dozer and an excavator. Both move on a predefined path and emit sounds to make the VTE more realistic. As explained before, the PEs (see Figure 7 in plan view) of the machines represent spheres with a radius of 12 meters.

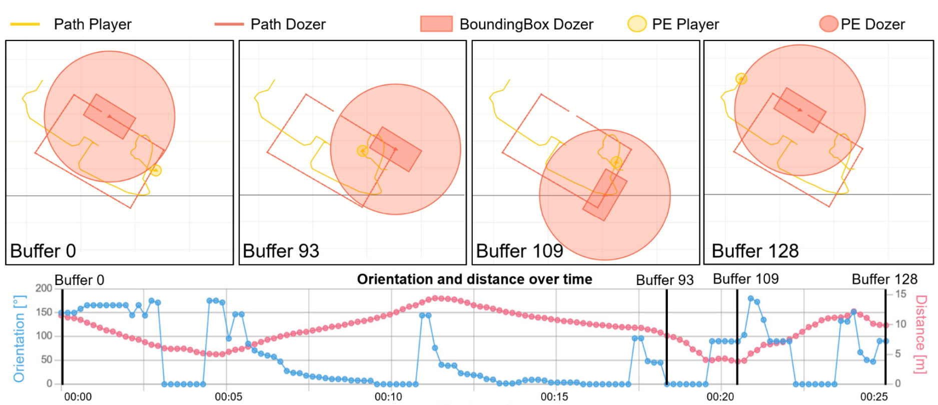

Figure 7: Four states of a close call related to a dynamic hazard (dozer).

Figure 7 also visualises the recorded close call. The event endures approximately 25 seconds and comprises 129 buffers (5 per second). The prototype visualises all states in an animation. In this paper, however, we can only illustrate a selection and interpret the results with the information in Figure 7. When the player approaches the dozer (Buffer 0), the orientation is between 150 and 180. A value of 180 entails that they directly approach each other, which increases the likelihood of a direct hit. The distance varies between 5 and 10 meters. The trajectories also imply that the player moved rather undetermined around the vehicle.

4.3 Data Collection and Analyses

The possibility to collect player-related data throughout the training experience allows for the automatic generation of personalised feedback based on the data collection and analyses described previously. As previously mentioned, such feedback improves the training experience. The implemented prototype utilises the proposed data model to analyse the collected data and visualises the results in a dashboard. Figure 8 illustrates what such feedback may contain. An overview of the training scene visualises the player's paths and all close calls. On the top right, the player receives an overall grade. The radar plot below shows the average of the five weights for all close calls, and the last chart indicates the near misses by hazard type. It was not in the scope of this work to investigate which data provides vital feedback to the workers. The figure below is just an example and future work needs to investigate how a grading scale may relate performances of individual workers.

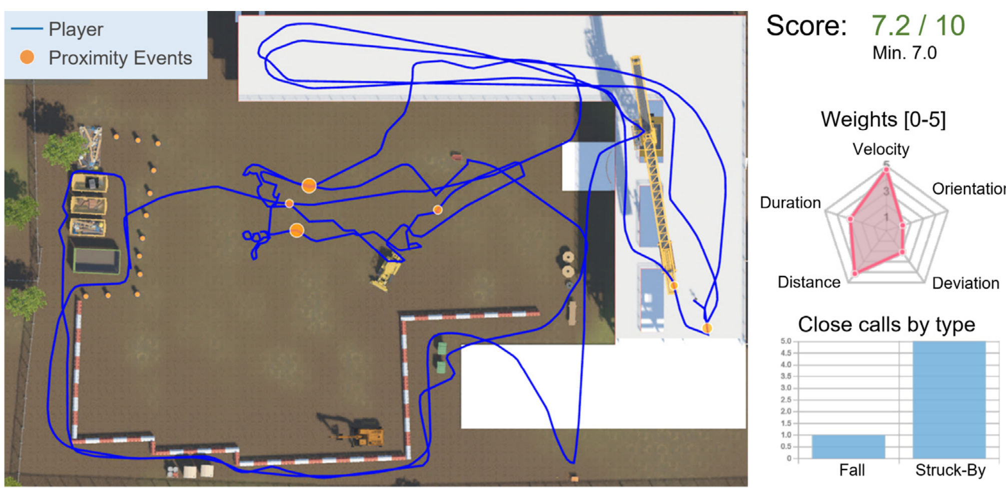

Figure 8: Personalised feedback for the player.

5. Conclusion

This paper proposed an ontology-based data model for automatically generating a VTE for construction safety. The VTEs allow safety officers to train workers in virtuality and assess their performance based on close calls. The data model was implemented in a prototype using IFC for the 4D BIM model, a graph-based database, and Unity for developing the training scene. The prototype was tested for a college building and a railway infrastructure project. The first results show that the data model allows for generating a training environment with little means. The data model also allows for data collection, which can further be utilised to provide personalised feedback. The tests highlighted that the DTC must provide high-quality data. For instance, every task must implement the necessary resources on a detailed level.

We conclude that the prototype proves the viability of the data model. However, additional case studies could be conducted to validate the proposed data model across various construction projects and hazards. This would help to establish the robustness and adaptability of the model and provide more extensive evidence of its effectiveness. Additionally, further research could investigate the impact of incorporating additional data sources such as real-time sensor data and worker behaviour data. This could enhance the accuracy and relevance of the VTE, leading to better safety outcomes. Lastly, the authors propose to further exploit the generated data of the training games. The data could support decision-makers in designing safe construction sites or train machine learning algorithms. Such an algorithm could predict hazards during construction based on the data from the virtuality.

Acknowledgements

The research presented in this paper has been funded by the European Union Horizon 2020 research and innovation program under grant agreement no. 958310 and 95398.

References

Ackoff, R. L. (1989). From data to wisdom. Journal of Applied Systems Analysis. 16, 1 (1989), 3–9.m data to wisdom. Journal of Applied Systems Analysis, 16(1).

Arcventure. (2023). IFC Importer. https://Assetstore.Unity.Com/Packages/Tools/Utilities/Ifc-Importer-162502.

Bernardos, S., Povedal-Villalon, M., Garcia-Castro, R., Giannakis, G., Tsalis, C., Charbi, T., Karalis, E., Teizer, J., Bosche, F., Karkanis, V., Falcioni, D., Moraitis, P., & Varga, J. (2022). D3.4-COGITO Data Model and Ontology Definition and Interoperability Design v3.

Bükrü, S., Wolf, M., Böhm, B., König, M., & Teizer, J. (2020). Augmented virtuality in construction safety education and training. EG-ICE 2020 Workshop on Intelligent Computing in Engineering, Proceedings.

Bureau of Labor Statistics. (2022). National census of fatal occupational injuries in 2021. https://www.bls.gov/news.release/pdf/cfoi.pdf

Golovina, O., Kazanci, C., Teizer, J., König, M. (2019a). "Using Serious Games in Virtual Reality for Automated Close Call and Contact Collision Analysis in Construction Safety." 36th International Symposium on Automation and Robotics in Construction, https://doi.org/10.22260/ISARC2019/0129.

Golovina, O., Perschewski, M., Teizer, J., & König, M. (2019b). Algorithm for quantitative analysis of close call events and personalised feedback in construction safety. Automation in Construction, 99. https://doi.org/10.1016/j.autcon.2018.11.014

Golovina, O., Teizer, J., Johansen, K. W., & König, M. (2021). Towards autonomous cloud-based close call data management for construction equipment safety. Automation in Construction, 132. https://doi.org/10.1016/j.autcon.2021.103962

Golovina, O., Teizer, J. (2022). "Serious game in Virtual Reality for Safe, Active, and Personalised Learning related to Pedestrian Workers Struck by Equipment and other Construction Hazards." 22nd Conference on Construction Application of Virtual Reality, 260-271, ISBN: 978-0-9927161-4-1.

Grady, B., Rumbaugh, J., & Jacobson, I. (2005). Unified Modeling Language User Guide. ResearchGate.

Harichandran, A., Johansen, K. W., Jacobsen, E. L., & Teizer, J. (2021). A Conceptual Framework for Construction Safety Training using Dynamic Virtual Reality Games and Digital Twins. Proceedings of the International Symposium on Automation and Robotics in Construction, 2021-November. https://doi.org/10.22260/isarc2021/0084

Hilfert, T., Teizer, J., König, M. (2016). "First Person Virtual Reality for Evaluation and Learning of Construction Site Safety." 33rd International Symposium on Automation and Robotics in Construction, Auburn, Alabama, USA, https://doi.org/10.22260/ISARC2016/0025.

IfcOpenShell. (2023, February 14). IfcConvert: An application for converting ifc geometry into several file formats. Https://Ifcopenshell.Sourceforge.Net/Ifcconvert.Html.

ISO 5006:2017. (2017). Earth-moving machinery - Operator's field of view - Test method and performance criteria.

Jacobsen, E. L., Solberg, A., Golovina, O., & Teizer, J. (2022). Active personalised construction safety training using run-time data collection in physical and virtual reality work environments. Construction Innovation, 22(3). https://doi.org/10.1108/CI-06-2021-0113

Johansen, K. W., Schultz, C., & Teizer, J. (2023). Hazard ontology and 4D benchmark model for facilitation of automated construction safety requirement analysis. Computer-Aided Civil and Infrastructure Engineering. https://doi.org/10.1111/mice.12988

Li, B., Schultz, C., Teizer, J., Golovina, O., & Melzner, J. (2022). Towards a unifying domain model of construction safety, health and well-being: SafeConDM. Advanced Engineering Informatics, 51. https://doi.org/10.1016/j.aei.2021.101487

Nykänen, M., Puro, V., Tiikkaja, M., Kannisto, H., Lantto, E., Simpura, F., Uusitalo, J., Lukander, K., Räsänen, T., Heikkilä, T., & Teperi, A. M. (2020). Implementing and evaluating novel safety training methods for construction sector workers: Results of a randomised controlled trial. Journal of Safety Research, 75. https://doi.org/10.1016/j.jsr.2020.09.015

Occupational Safety and Health Administration. (2016). Recommended Practices for Safety and Health Programs in Construction. www.osha.gov

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualising, measuring, and improving the capacity of classroom interactions. In Handbook of Research on Student Engagement. https://doi.org/10.1007/978-1-4614-2018-7_17

Sacks, R., Brilakis, I., Pikas, E., Xie, H. S., & Girolami, M. (2020). Construction with digital twin information systems. Data-Centric Engineering, 1(6). https://doi.org/10.1017/dce.2020.16

Sacks, R., Perlman, A., & Barak, R. (2013). Construction safety training using immersive virtual reality. Construction Management and Economics, 31(9). https://doi.org/10.1080/01446193.2013.828844

Schlenger J, Yeung T, Vilgertshofer S, Martinez J, Sacks R, & Borrmann A. (2022). A Comprehensive Data Schema for Digital Twin Construction. 29 Th International Workshop on Intelligent Computing in Engineering. https://publications.cms.bgu.tum.de/2022_Schlenger_EGICE.pdf

Stanney, K., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., Weech, S., & Fulvio, J. M. (2020). Identifying Causes of and Solutions for Cybersickness in Immersive Technology: Reformulation of a Research and Development Agenda. International Journal of Human-Computer Interaction, 36(19). https://doi.org/10.1080/10447318.2020.1828535

Teizer, J., Cheng, T., & Fang, Y. (2013). Location tracking and data visualisation technology to advance construction ironworkers' education and training in safety and productivity. Automation in Construction, 35, https://doi.org/10.1016/j.autcon.2013.03.004

Teizer, J., Johansen, K. W., & Schultz, C. (2022). The Concept of Digital Twin for Construction Safety. Construction Research Congress, https://doi.org/10.1061/9780784483961.121

The Center for Construction Research and Training. (2022). Fatal and Nonfatal Injuries in the Construction Industry. https://www.bls.gov/iif/oshwc/cfoi/cfoi_rates_2020hb.xlsx.

Wolf., M., Teizer, J., Wolf, B., Bükrü, S., Solberg, A. (2022). "Investigating hazard recognition in augmented virtuality for personalised feedback in construction safety education and training." Advanced Engineering Informatics, 51, 101469, https://doi.org/10.1016/j.aei.2021.101469.

Wolf, M., Teizer, J. (2022). "Serious Game Design Framework to Teach Hazard Awareness and Avoidance on Construction Sites using Pre-defined Hazard Definitions." 22nd Conference on Construction Application of Virtual Reality (CONVR), 348-358, ISBN: 978-0-9927161-4-1.

Zheng, Y., Törmä, S., & Seppänen, O. (2021). A shared ontology suite for digital construction workflow. Automation in Construction, 132. https://doi.org/10.1016/j.autcon.2021.103930